DOCUMENT RESUME

ED 320 929 TM 015 148

AUTHOR Ferrell, Barbara G.
TITLE Exstatix: Expandable Statistical Analysis System for the Macintosh. A Software Review.
PUB DATE Apr 90
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 16-20, 1990).
PUB TYPE Book/Product Reviews (072) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Analysis of Variance; Computer Software Evaluation; *Educational Research; Graphs; Microcomputers; Regression (Statistics); Research Reports; *Research Tools; *Statistical Analysis; Tables (Data)
IDENTIFIERS *Exstatix

ABSTRACT

The Exstatix statistical analysis software package by K. C. Killion for use with Macintosh computers is evaluated. In evaluating the package, the framework developed by C. J. Ansorge et al. (1986) was used. This framework encompasses features such as transportability of files, compatibility of files with other Macintosh software, and ability to merge results with word processing and desk top publishing applications. The review is from the perspective of the educational researcher who wishes to use the package for uncomplicated, quick-and-dirty analysis of relatively small univariate data sets and export the output to reports. Specific features covered in the evaluation include documentation, ease of use, running programs, program output, accuracy, and other features. Strengths of Exstatix are discussed in the areas of statistical decision making, assumption testing, and regression analysis. Weaknesses of the packages include problems with importing and exporting functions, report generation, analysis of variance applications, and add-ons. Sample output of tables and graphs are provided. (TJH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

BARBARA G. FERRELL

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

ED320929

Exstatix™ Version 1.0.4

by Kevin C. Killion

Select Micro Systems, Inc.
40 Triangle Center - Suite 211
Yorktown Heights, NY 10598 (914) 245-4670

©Copyright, 1989, Stone House Systems. All Rights Reserved.

Barbara G. Ferrell, Ph.D.
University of Texas Medical Branch

Paper Presented at AERA, Boston, MA
April 17, 1990

TM015148

BEST COPY AVAILABLE

Exstatix™: Expandable Statistical Analysis System for the Macintosh™

A Software Review

Barbara G. Ferrell, Ph.D.
The University of Texas Medical Branch at Galveston

Statistical software packages for the Macintosh™ have proliferated over the last three years. These packages have been described as a "bumper crop" (Levine, 1990) and a "powerful force in statistics" (Seiter, 1989) in the Macintosh-related magazines. The focus of these reviews, however, is usually business application, and the articles are written by programmers or other computer types and not statisticians. The viewpoint of the researcher who will utilize the program in data analysis is always missing.

Because statistical software does not find a ready market in the general population of Mac users, most computer stores carry them only if they have a business focus. Potential users have only the product reviews on which to rely. The criteria used to evaluate a statistics package for a researcher is different from that for business. The purpose of this paper is to review Exstatix™, produced by Select Micro Systems, Inc., from the perspective of an educational researcher.

Evaluation Paradigm

Other statistical software packages for the Macintosh have been evaluated by this reviewer at previous sessions of AERA (Ferrell, 1989; Ferrell, 1988). Exstatix™ is similar to those previously reviewed in that it might be viewed as a "personal" package; software that is inexpensive enough than an individual would be willing to buy it from his/her own pocket or recommend it to a student to use to analyze thesis data. Exstatix™ is the newest of the software in this category.

In evaluating Exstatix™, the framework developed by Ansorge, Wise and Plake (1986) was used, but other factors (Hamer, 1981) such as transportablitiy of files, compability of files with other Mac software, ability to merge results with word processing and desk top publishing applications, etc., were considered. The review is from the perspective of the educational researcher who wishes to use the package for uncomplicated, quick-and-dirty analysis of relatively small univariate data sets and export the output into reports.

General Information

Exstatix™
Version 1.0.4

Vendor:

Select Micro Systems, Inc.
40 Triangle Center, Suite 211
Yorktown Heights, NY 10598
(914) 245-4670

Price:	$349 Suggested Retail Educational Discounts and "Student Pack" Available
System Requirements:	Macintosh 512KE, Plus, SE. Color supported on Macintosh II. System 4.1 or later. Supports Imagewriter™, Laserwriter™ and other printers.
Statistics:	Descriptive, Crosstabs, Autocorrelation, Correlation, Regression, t-test, N-way ANOVA, Scheffe, Non-parametrics, Time Series, Transformations, Box Plots, Scatter Plots, Bar Chart, Pie Chart, 3-D Scatter & Bar Charts.

Software Usability

Documentation

Exstatix™ comes with a manual that assumes that the user knows something about the Macintosh™. Mac novices are referred to the computer manual and no teaching of Mac's standard features or "mouse ability" are included.

The seasoned statistician is directed to an overview of features designed for "those experienced in the use of Statistics AND the Macintosh" (Wygonik, Killion & Scott, 1988). The novice statistician is urged to use the extensive tutorial which comprises the major portion of the manual. The tutorial is more than the usual how-to-do-it. It is in some sense a mini-applied statistics course, providing the user with guidance as to when to apply the tests. It is divided into sections dealing with analyses for one variable, two variables, and three or more variables.

The manual is clearly written and complete. Appendices include sample problems, some sections on statistical decision making and formulae for the analyses.

Ease of Use

Exstatix™ is menu driven and easy to use. It is a single disk program, so no juggling is necessary if the user has only one drive, except when data are to be stored on a separate disk. The program has taken the features of the Macintosh™ into account in its development and did not just clone PC statistical package features. For example, variables can be viewed as icons on the screen and the user clicks on the icons to select variables for analysis.

No on-line help screens are available, but error diagnostics are good. Warnings are issued when the assumptions of a technique are violated, and they are written in non-statistical language. For example, to alert the researcher to small expected cell size in a Chi Square analysis, the following warning is issued:

Warning!
Of 4 cells,
2 have expected values of less than five

Additionally, analyses based on improper use of the data will not be performed. When asked for a correlation matrix using all of the variables, alpha variables were excluded from the analysis.

Running Programs

To test the running of Exstatix™, the data set from previous evaluations of statistical software was used. It was in the attempt to import the data set from another statistics package that a limitation of Exstatix™ that has been cited by other reviewers (MacWorld, May, 1989) was demonstrated.

While the Exstatix™ literature states that Exstatix™ files and ASCII text files can be imported or exported, the manual suggests that data files from other Mac applications can be imported if they have been saved in text format. Data files from other statistical programs, the most likely candidates for import, are not usually in text file. Importing was accomplished, however, by copying the file to the clipboard and pasting it into Exstatix™. Variable labels were lost during the process and had to be retyped.

In terms of data analysis, Exstatix™ has a strength in its regression analysis subroutine. Not only does the program permit standard multiple regression analysis, but stepwise, both backward and forward, and subset analyses are available. The linear regression program provides for crossvalidation and a variety of optional tests and output including the regression equation, and nonlinear (log, exponential) and polynomial analyses are available. The regression analysis subroutine in Exstatix™ is the most comprehensive of statistics packages in its price range. (See sample of output in the Appendix.)

Just as the regression subroutine is a strength of Exstatix™, the ANOVA subroutine is a weakness. Other reviews have referred to its ANOVA procedures as "minimal and hard to specify" (Levine, 1990). In attempting to run the ANOVA procedures, real difficulty was encountered in the way the algorithm is set up to handle the data. Two major problems make this subroutine of very little use in most data sets utilized in educational research. First, Exstatix™ ANOVA requires that the data be set up with each level of a factor represented by a column for at least one factor in the data set. If this method is followed, specifying a one-way ANOVA will enable the researcher to obtain the summary table for either an equal or an unequal cell size design. Specifying a two-way ANOVA will give the researcher a repeated measures analysis, because the rows are considered one independent variable while the columns are the second. With the N-way ANOVA algorithm, equal cell sizes are required and the manual specifies that the researcher must set up the data "with each case representing a complete additional experiment" (Wygonik, Killion, & Scott, 1988, p. 107) and the design must be completely balanced. For the most part, this setup makes the ANOVA subroutine entirely unusable to the researcher who rarely has equal group sizes. In addition, using the N-way routine would require making a separate data set for that analysis only, an awkward method that is not necessary with other statisitics packages of the same type as Exstatix™.

A good feature of Exstatix™ from the viewpoint of the statistics teacher or the statistical consultant is that it is the only program that encorporates research design and statistical decision making and takes the assumptions of the tests into account. Variables can be designated as independent or dependent, making the user at least think about design. While the usual p values also are printed, the significance level selected remains set until changed by the researcher. Alpha level is preset by the user prior to running the test and statistical decisions based on that level are printed. These two features plus the warnings which are issued when problems such as multicollinearity or small expected

cell size occur make this the only program which pays even lip service to the proper use of statistical findings.

Program Output

Just as the import of data to Exstatix™ was problemmatic, so was the export of tables produced. When pasted into both desk top publishing and word processing programs, tabs were lost and the font was changed to the default New York font 12 point. Tables, however, are fully editable, so that they need not be pasted to a paint program for adding titles or other information before placing in documents. However, if tables need to be edited once placed into other documents, it is questionable as to whether or not this represents a savings in time. The exported tables were definately not of a quality which the reviewer would consider "document ready." Examples of program output are included in the appendix.

A variety of graphics displays are available on Exstatix™. Scatter and box plots may be displayed, and both two and three dimensional bar and scatter plots are possible. The 3-D options may be used to rotate the scatterplot, a simple version of some of the exploratory data analysis programs available for the Mac. Quality of output is adequate. Examples of graphics displays are in the appendix.

Accuracy

Exstatix™ uses Standard Apple Numeric Environment (SANE) arithmetic and provisional means algorithms for accuracy. No information regarding the number of significant digits computed was provided in the manual. Exstatix™ computes a variety of statistics during idle time, thus speeding up subsequent analyses. It can handle up to 100 variables with as many as 1000 cases per variable.

Other Features

Exstatix™ has been marketed as "the expandable statistical analysis system" based on its capability to add new functions and commands in the form of "custom packets." Two types of packets may be added, function packets which enable the user to add to the Functions menu, and command packets which place new commands on the menu. This feature would enable the user to add new procedures or to customize those already availble for a special purpose. The ability to expand Exstatix™, however, requires knowledge of Pascal or C. The software developers do not intend to add packets, but instead will rely on outside sources for these. While it is possible that good packets may be developed to go along with the basic program, it seems that those with advanced knowledge of Pascal or C probably would not select this program in the first place, but would work with one of the higher end statistical packages which are available.

Summary

Based on the review of Exstatix™ the following is noted:

Strengths	Weaknesses
statistical decision making	importing/exporting
assumption testing	reports
regression	ANOVA
	addons

RATING*

DOCUMENTATION	ΣΣΣ
EASE OF USE	ΣΣΣ
RUNNING PROGRAMS	ΣΣ
PROGRAM OUTPUT	Σ
OTHER	Σ
OVERALL RATING	ΣΣ

Rating based on a maximum of ΣΣΣΣΣ.

References

Ansorge, C.J., Wise, S.L. & Plake, B.S. (1986). Evaluating the quality ot microcomputer software. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Ferrell, B.G. (1989). Statistical software package for the Macintosh: Different software for different users. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA.

Ferrell, B.G. (1988). Two statistical software packages for the Mac. Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Hamer, R.M. (1981). Papers that evaluate computer programs. American Statiscian, 35(4), 264.

Levine, M. (April, 1990). Business statistics. MacUser, 115-134.

MacWorld (May, 1989). Statistical software: Exstatix. MacWorld, 189-191.

Seiter, C. (April, 1989). Picturing statistics: Statistics programs that help you analyze numbers with graphics. MacWorld, 138-143.

Appendix
Sample Output

Descriptive*

*** Tables have been edited for ease of interpretation. For actual output see display.**

Science Project Sample Output from Exstatix™

Summary statistics for all numeric variables:

	Mean	St. Deviation	St. Error	Median	Range
Case Number	18.50	10.54	1.76	18.50	35.00
Boating Experience	21.86	11.17	1.86	18.50	44.00
Seasickness Exper...	22.50	22.88	3.81	5.50	113.00
GEFT	13.67	5.43	0.91	16.00	17.00
Age In Years	44.06	13.60	2.27	45.00	50.00
Index	1.79	3.84	0.64	0.71	22.60

Basic statistics for GEFT:

Active cases:	36
Missing values:	0
Mean:	13.6667
Sum:	492.0000
Sum of squares:	7756.0000
Variance:	29.4857
Std deviation:	5.4301
Std error:	0.9050
Second moment:	28.6667
Third moment:	-207.1296
Fourth moment:	2725.1481
Skewness:	-1.3495
Kurtosis (norm=3):	3.3162
Median:	16.0000
Minimum:	1.0000
Maximum:	18.0000
Range:	17.0000

Frequency Distributions and Crosstabs

Frequency table for Sex:

Active cases included: 36

	Count	% Total
Female	16	44.4
Male	20	55.6

Crosstab of:

- **Field Ind/Dep**
- **Sex**

Cases included: 36

	Field Ind/Dep		
	Dependent	Independent	Total
Sex			
Female			
Count	2	14	16
% across	12.5	87.5	100.0
% down	28.6	48.3	44.4
% total	5.6	38.9	44.4
index	64	109	100
expected	3.1	12.9	16.0
deviation	-1.1	1.1	0.0
Male			
Count	5	15	20
% across	25.0	75.0	100.0
% down	71.4	51.7	55.6
% total	13.9	41.7	55.6
index	129	93	100
expected	3.9	16.1	20.0
deviation	1.1	-1.1	0.0
Total			
Count	7	29	36
% across	19.4	80.6	100.0
% down	100.0	100.0	100.0
% total	19.4	80.6	100.0
index	100	100	100
expected	7.0	29.0	36.0
deviation	0.0	0.0	0.0

Degrees of freedom: 1
Chi-square: 0.887
Yates corrected chi-square: 0.268
Association (phi): 0.025
Cramer's V: 0.157
Contingency coefficient: 0.155

Warning! Of 4 cells, 2 have expected values of less than five

Correlation and Regression

Table of correlations:

	Case Number	Boating Expe...	Seasickness ...	GEFT
Case Number	1.000	-0.455Δ	0.192	-0.178
Boating Experience	-0.455Δ	1.000	-0.178	0.195
Seasickness Exper...	0.192	-0.178	1.000	-0.355Δ
GEFT	-0.178	0.195	-0.355Δ	1.000
Age In Years	0.221	-0.010	0.149	-0.362Δ
Index	0.304	-0.398Δ	0.846Δ	-0.305

	Age In Years	Index
Case Number	0.221	0.304
Boating Experience	-0.010	-0.398Δ
Seasickness Exper...	0.149	0.846Δ
GEFT	-0.362Δ	-0.305
Age In Years	1.000	0.268
Index	0.268	1.000

Δ denotes correlations significantly different from zero, at the 95% level.

Table of covariances:

	Case Number	Boating Expe...	Seasickness ...	GEFT
Case Number	111.000	-53.586	46.229	-10.171
Boating Experience	-53.586	124.866	-45.500	11.810
Seasickness Exper...	46.229	-45.500	523.571	-44.057
GEFT	-10.171	11.810	-44.057	29.486
Age In Years	31.600	-1.478	46.400	-26.724
Index	12.282	-17.050	74.260	-6.357

	Age In Years	Index
Case Number	31.600	12.282
Boating Experience	-1.478	-17.050
Seasickness Exper...	46.400	74.260
GEFT	-26.724	-6.357
Age In Years	184.968	13.987
Index	13.987	14.731

Table of partial correlations:

(Each entry shows correlation of variables of that row and column, after partialling out the effects of other variables in the table.)

	Case Number	Boating Expe...	Seasickness ...	GEFT
Case Number	1.000	-0.389	0.040	0.013
Boating Experience	-0.389	1.000	0.377	0.219
Seasickness Exper...	0.040	0.377	1.000	-0.316
GEFT	0.013	0.296	-0.295	1.000
Age In Years	0.203	0.296	-0.295	-0.375
Index	0.009	-0.471	0.861	0.192

	Age In Years	Index
Case Number	0.203	0.009
Boating Experience	0.296	-0.471
Seasickness Exper...	-0.295	0.861
GEFT	-0.375	0.192
Age In Years	1.000	0.349
Index	0.349	1.000

Compare ranks of:
Age In Years
GEFT

Test used: Spearman Rank Correlation
Degrees of freedom: 34
Spearman R: -0.4531
T-statistic: -2.9635

Test used: Kendall Rank Correlation
S: -203
Kendall tau: -0.3456

Regression of: Seasickness Experience

Independent variables:
Boating Experience
GEFT
Sex
Numeric equivalents used for alpha items:
1 = Female
2 = Male
Age In Years
Index

	Multiple-R	R-square
Unadjusted:	0.9021	0.8138
Adjusted:	0.8847	0.7828

Sample size = 36
Standard Error of Estimate = 10.6647
F = 26.224
p = 0.0000

Is the F-value statistically significant?
Required critical value of F for 95% level of significance is 2.679.
The achieved F-value IS significant at the 95% level.

	Regression Coefficient	Standard Error of Coefficient	T Value	Beta Weight
Boating Experie...	0.493270	0.179057	2.7548 (p= 0.0099)	0.240890
GEFT	-0.744881	0.368961	-2.0189 (p= 0.0525)	-0.176769
Sex	-8.790018	3.717654	-2.3644 (p= 0.0247)	-0.193594
Age In Years	-0.186420	0.149927	-1.2434 (p= 0.2233)	-0.110803
Index	5.504803	0.541939	10.1576 (p= 0.0000)	0.923356
Intercept	33.943135			

Seasickness Experience = 33.94313456 + 0.49326966 * Boating Experience - 0.74488146 * GEFT - 8.79001776 * Sex - 0.18641991 * Age In Years + 5.5048033 * Index

Tests on residuals:
Available cases: 36
Cases with error > 0: 16
Cases with error < 0: 20

Standard error of estimate (standard deviation): 10.6647
Mean absolute error: 7.60096893
Greatest error: 23.32015644

Durbin-Watson: 2.294
Von Neumann ratio: 2.360

ANOVA of current regression:

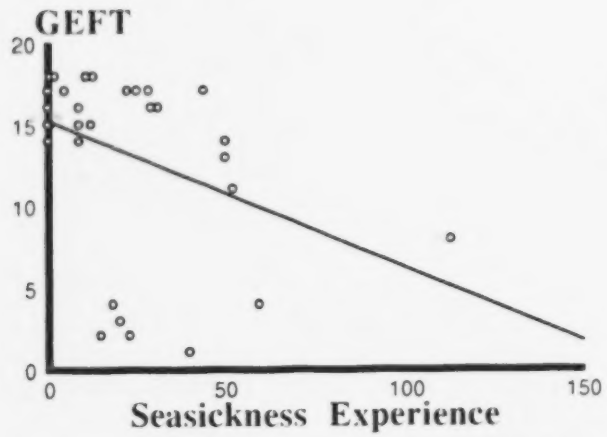
Source	Sum of Squares	Degrees of Freedom	Mean Squares
Regression:	14912.92451	5	2982.58490
Residuals:	3412.07549	30	113.73585
Total:	18325.00000	35	523.57143

T-Test

(Independent values; variances assumed equal)
Degrees of freedom: 62
T-statistic: 4.8488

Is the DIFFERENCE statistically significant?
Required t value for 95% level of significance is 1.999
The difference IS significant at the 95% level.
Is the DIRECTION of the difference statistically significant?
Required t value for 95% level of significance is 1.670
The direction of the difference IS significant at the 95% level.

Scatterplot



3-D Bar Chart